CSE 559A: Computer Vision

[credit: danjodon.deviantart.com]

Fall 2017: T-R: 11:30-1pm @ Lopata 101

Instructor: Ayan Chakrabarti (ayan@wustl.edu).
Staff: Abby Stylianou (abby@wustl.edu), Jarett Gross (jarett@wustl.edu)

http://www.cse.wustl.edu/~ayan/courses/cse559a/

Oct 26, 2017

GENERAL

- Problem Set 3 Due Thursday.
- Project Proposals Deadline Extended to 11:59 PM Tuesday Oct 31st.
    - Submitted through blackboard.
    - 2-3 Paragraphs. Can be PDF / Text File / Put directly in the text box.

- Push back PSET 4/5 by two days each.

- PSET 4 Will be Posted Tuesday (will include stuff we do next week).
    - Will be due two weeks from then.

- PSET 5 will be pushed back. Now due Thu after TG break instead of Tue.

New schedule is online.

STEREO ROUNDUP

Last time: "message" passing / augmented cost

$$\bar{C}[x,d] = C[x,d] + \min_{d'} \bar{C}[x-1,d'] + \lambda S(d,d')$$

- Consider the case when $S(d,d')$:
  - 0 if $d = d'$
  - P_1 if $|d - d'| = 1$
  - P_2 otherwise.
- Can we do this efficiently ?
  - Need to go through each line sequentially.
  - But can go through all lines in parallel.
  - But what about d ? Do we need to do minimization for every d independently ?

STEREO ROUNDUP

$$\bar{C}[x, d] = C[x, d] + \min_{d'} \bar{C}[x - 1, d'] + \lambda S(d, d')$$

- Note: It doesn't matter if we add / subtract constants to all d's:
  - $C[x, d]$ with $C[x, d] + C_0[x]$
  - $\bar{C}[x, d]$ with $\bar{C}[x, d] + C_0[x]$

Why not ?

- Because the minimization will always be over d. You are never comparing $C[x_1, d_1]$ with $C[x_2, d_2]$.

STEREO ROUNDUP

$$\bar{C}[x,d] = C[x,d] + \min_{d'} \tilde{C}[x-1,d'] + S(d,d')$$

$$S(d,d') = \begin{cases} 0 \text{ if } d = d' \\ P_1 \text{ if } |d-d'| = 1 \\ P_2 \text{ otherwise} \end{cases}$$

- Step 1 (Simplify): Replace $\bar{C}[x-1,d']$ with $\tilde{C}[x-1,d'] = \bar{C}[x-1,d'] - \min_{d''} \bar{C}[x-1,d'']$

The MAXIMUM value for $\min_{d'} \tilde{C}[x-1,d'] + S(d,d')$ is P_2.

- Step 2: This means that for every value of d, we just need to consider four values.

- $\min_{d'} \tilde{C}[x-1,d'] + S(d,d')$ is the min of
  - P_2
  - $\tilde{C}[x-1,d-1] + P_1$
  - $\tilde{C}[x-1,d+1] + P_1$
  - $\tilde{C}[x-1,d]$

STEREO ROUNDUP

$$\bar{C}[x,d] = C[x,d] + \min_{d'} \tilde{C}[x-1,d'] + S(d,d')$$

$$S(d,d') = \begin{cases} 0 \text{ if } d = d' \\ P_1 \text{ if } |d-d'| = 1 \\ P_2 \text{ otherwise} \end{cases}$$

- $\min_{d'} \tilde{C}[x-1,d'] + S(d,d')$ is the min of
  - P_2
  - $\tilde{C}[x-1,d-1] + P_1$
  - $\tilde{C}[x-1,d+1] + P_1$
  - $\tilde{C}[x-1,d]$

Can do this in parallel with matrix operations for all d and all lines.

Full algorithm in paper:
Hirschmueller, Stereo Processing by Semi-Global Matching and Mutual Information, PAMI 2008.

STEREO ROUNDUP

SGM Algorithm Averages along four directions:

$$\bar{C}_{lr}[n,d] = C[n,d] + \min_{d'} \bar{C}_{lr}[n - [1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{rl}[n,d] = C[n,d] + \min_{d'} \bar{C}_{rl}[n + [1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{du}[n,d] = C[n,d] + \min_{d'} \bar{C}_{du}[n - [0,1]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{ud}[n,d] = C[n,d] + \min_{d'} \bar{C}_{ud}[n + [0,1]^T, d'] + \lambda S(d,d')$$

$$d[n] = \arg\min_{d} \bar{C}_{lr}[n,d] + \bar{C}_{rl}[n,d] + \bar{C}_{ud}[n,d] + \bar{C}_{du}[n,d]$$

Bur $\bar{C}_{lr}$ is still smoothing the original cost.

STEREO ROUNDUP

SGM Algorithm Averages along four directions:

$$\bar{C}_{lr}[n,d] = (C[n,d] + \bar{C}_{rl}[n,d] + \bar{C}_{ud}[n,d] + \bar{C}_{du}[n,d]) + \min_{d'} \bar{C}_{lr}[n - [1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{rl}[n,d] = (C[n,d] + \bar{C}_{lr}[n,d] + \bar{C}_{ud}[n,d] + \bar{C}_{du}[n,d]) + \min_{d'} \bar{C}_{rl}[n - [1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{du}[n,d] = (C[n,d] + \bar{C}_{lr}[n,d] + \bar{C}_{rl}[n,d] + \bar{C}_{ud}[n,d]) + \min_{d'} \bar{C}_{du}[n - [1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{ud}[n,d] = (C[n,d] + \bar{C}_{lr}[n,d] + \bar{C}_{rl}[n,d] + \bar{C}_{du}[n,d]) + \min_{d'} \bar{C}_{ud}[n - [1,0]^T, d'] + \lambda S(d,d')$$

Wouldn't this be better ?

Why not this ?

Because this is a circular definition.

STEREO ROUNDUP

Loopy Belief Propagation (one version)

$$\bar{C}_{lr}^{t+1}[n,d] = (C[n,d] + \bar{C}_{rl}^{t}[n,d] + \bar{C}_{ud}^{t}[n,d] + \bar{C}_{du}^{t}[n,d]) + \min_{d'} \bar{C}_{lr}^{t+1}[n-[1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{rl}^{t+1}[n,d] = (C[n,d] + \bar{C}_{lr}^{t}[n,d] + \bar{C}_{ud}^{t}[n,d] + \bar{C}_{du}^{t}[n,d]) + \min_{d'} \bar{C}_{rl}^{t+1}[n-[1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{du}^{t+1}[n,d] = (C[n,d] + \bar{C}_{lr}^{t}[n,d] + \bar{C}_{rl}^{t}[n,d] + \bar{C}_{ud}^{t}[n,d]) + \min_{d'} \bar{C}_{du}^{t+1}[n-[1,0]^T, d'] + \lambda S(d,d')$$
$$\bar{C}_{ud}^{t+1}[n,d] = (C[n,d] + \bar{C}_{lr}^{t}[n,d] + \bar{C}_{rl}^{t}[n,d] + \bar{C}_{du}^{t}[n,d]) + \min_{d'} \bar{C}_{ud}^{t+1}[n-[1,0]^T, d'] + \lambda S(d,d')$$

Do this iteratively

More generally, at time step t, pass a message from node n to n', based on all messages n has at that time, except for the message from n'.

Read more:
- Yedidia, Freeman, Weiss, "Understanding belief propagation and its generalizations," IJCAI 2001 (Distinguished Paper)
- Tappen & Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters", ICCV 2003.

STEREO ROUNDUP

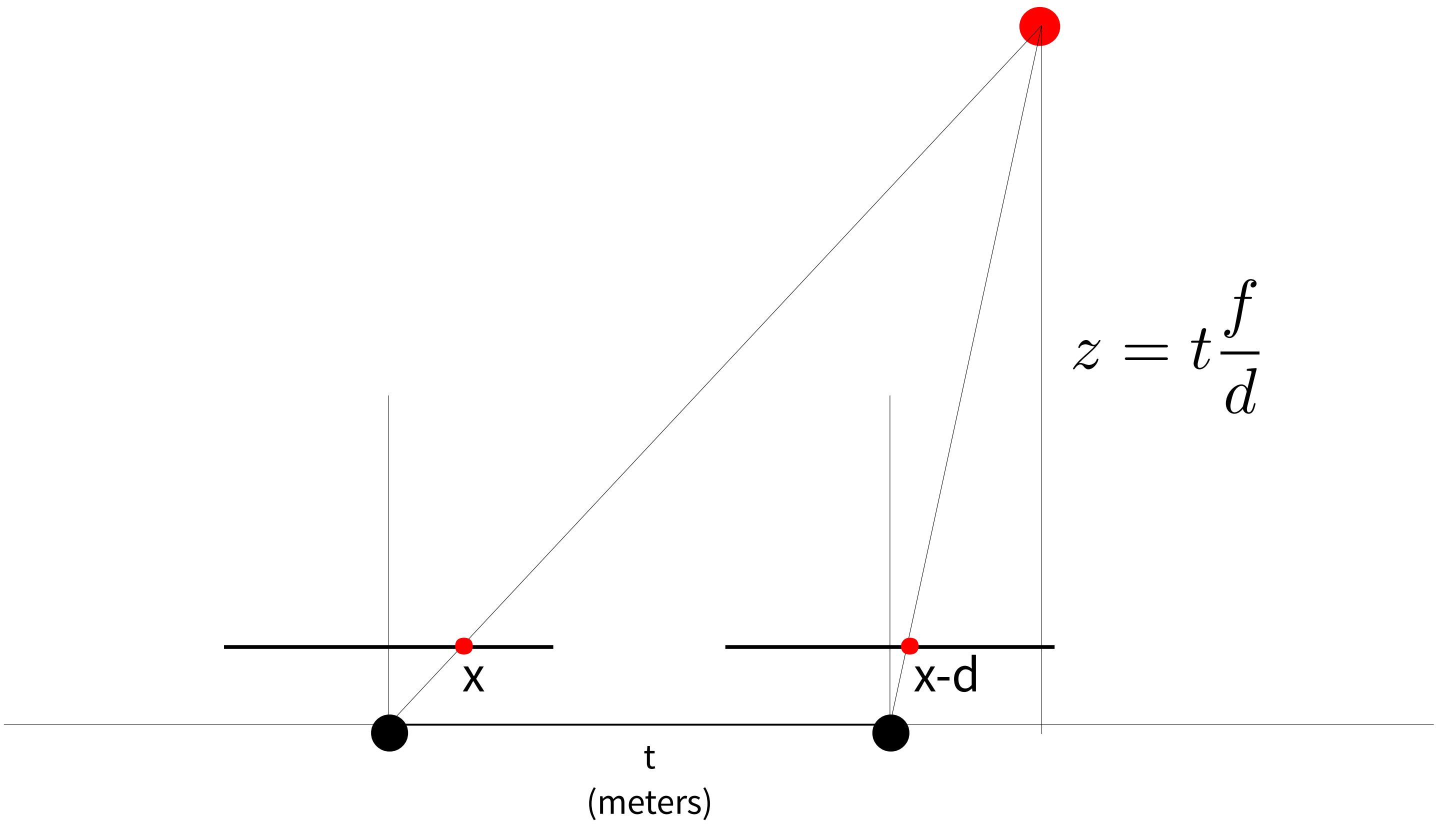
- Other methods for discrete minimzation---based on "Graph Cuts".
- SGM / Loopy BP: Generalize that there is an exact solution for a chain.
- Graph Cuts (with expansions / swaps): Generalize that there is an exact solution if only two values of d.

D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. IJCV 2002.

http://vision.middlebury.edu/stereo/

STEREO ROUNDUP

Disparity to Depth



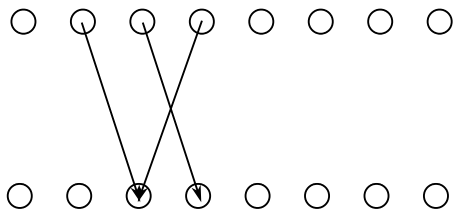
$$z = t\frac{f}{d}$$

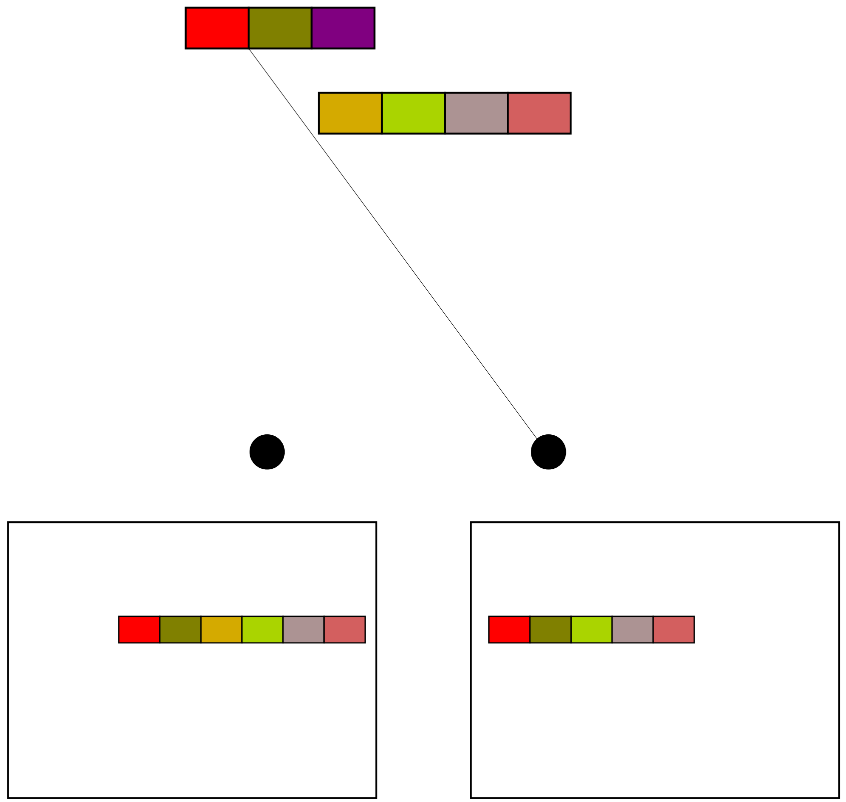
STEREO ROUNDUP

Uniqueness & Occlussions

Can encode this in the "smoothness cost"
along horizontal edges, putting an infinite cost
for disparity to right being higher than disparity to left,
and with additional "labels" for occlusions.

Pixels on Epipolar Lines





Can currently match arbitrarily

Some pixels in left image won't exist in right image.

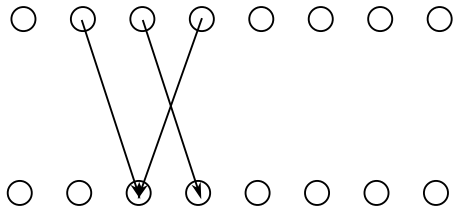
But, if a is to the left of b in the left image, and both appear in the right image, then a will still be to the left of b in the right image.

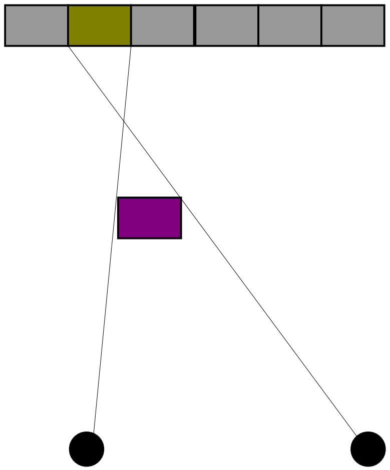
STEREO ROUNDUP

Uniqueness & Occlussions

Can encode this in the "smoothness cost"
along horizontal edges, putting an infinite cost
for disparity to right being higher than disparity to left,
and with additional "labels" for occlusions.

Pixels on Epipolar Lines





Can currently match arbitrarily

Some pixels in left image won't exist in right image.

But, if a is to the left of b in the left image, and both appear in the right image, then a will still be to the left of b in the right image.

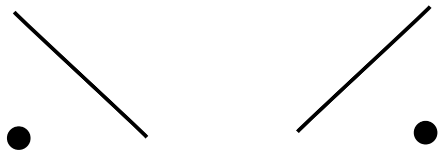
Doesn't always hold !

Also, often too complicated.

Just do matching in left & right,
check for consistency, remove
inconsistent matches, and do post-processing.

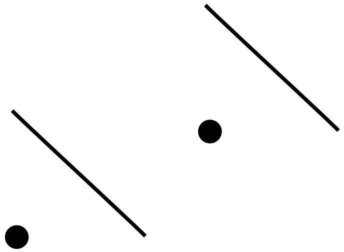
STEREO ROUNDUP

Rectified Stereo



STEREO ROUNDUP

Rectified Stereo



Un-rectified Stereo / Epipolar Flow

Yamaguchi, McAllester, Urtasun, "Robust monocular epipolar flow estimatin," CVPR 2013

Each point lies on some arbitrary line

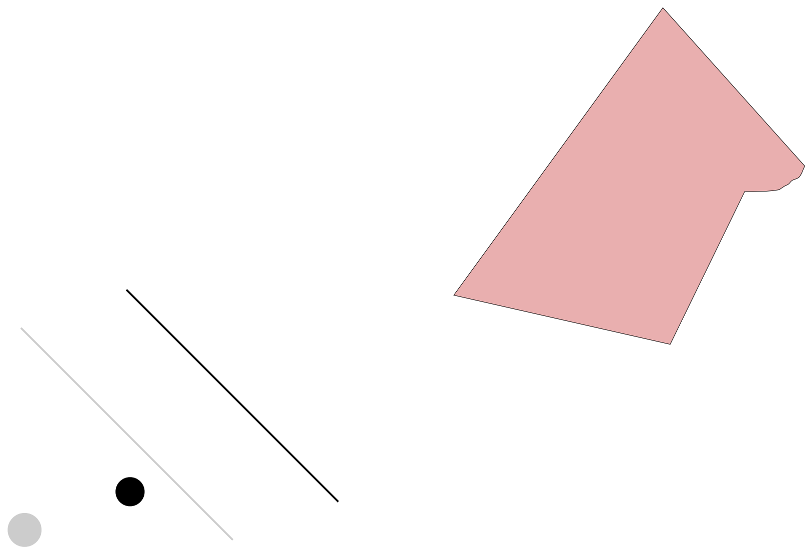
Search on this line, but how do you smooth ?

L[x,y] matches to R[x+u,y+v]

$$[u, v]^T = \underbrace{t_0(x, y)}_{\text{Displacement vector for a point at infinity}} + \overbrace{d}^{\text{Positive scalar inversely proportional to depth (smooth this)}} \underbrace{e_0(x, y)}_{\text{Unit vector in direction of epipolar line.}}$$

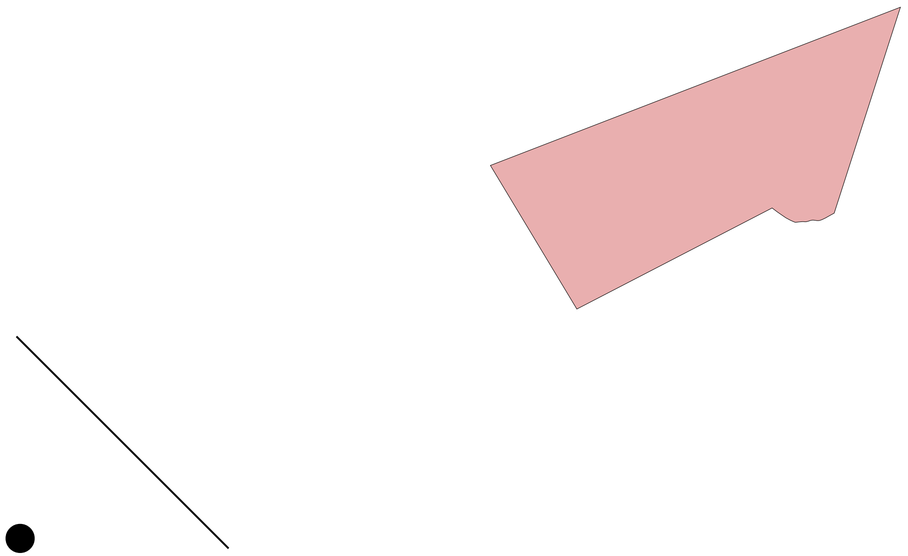
STEREO ROUNDUP

Epipolar Geometry Defines ***Relative*** Rigid Motion



STEREO ROUNDUP

Epipolar Geometry Defines ***Relative*** Rigid Motion



$$p_1^T F p_2 = 0$$

As long as all points that we're considering move rigidly between two images, i.e., their relative distances to each other don't change,

then the correspondences are related by epipolar geometry.

STEREO ROUNDUP

Stereo is too hard

Wouldn't it be great if we lived in a world where
everything was dark except for a single point on each horizontal line ?

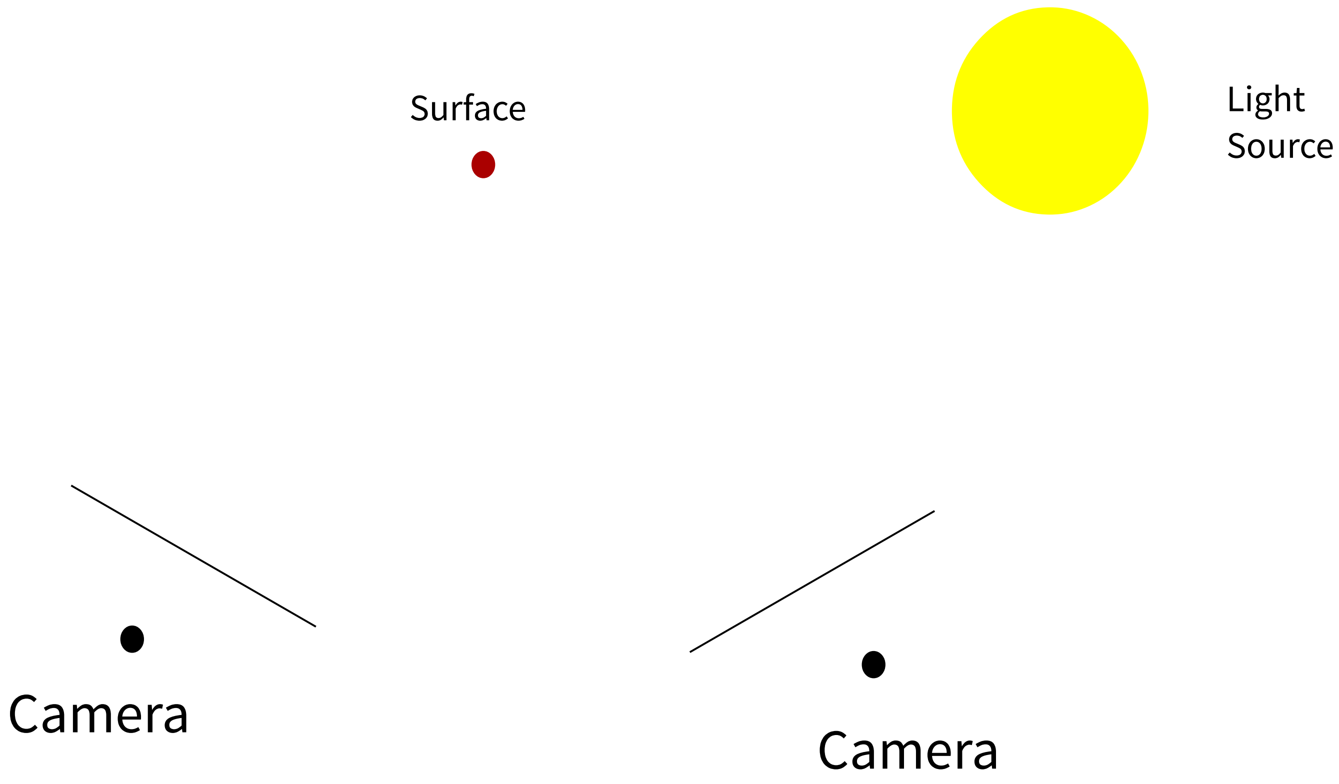
STEREO ROUNDUP

Stereo is too hard

Wouldn't it be great if we lived in a world where
everything was dark except for a single point on each horizontal line ?

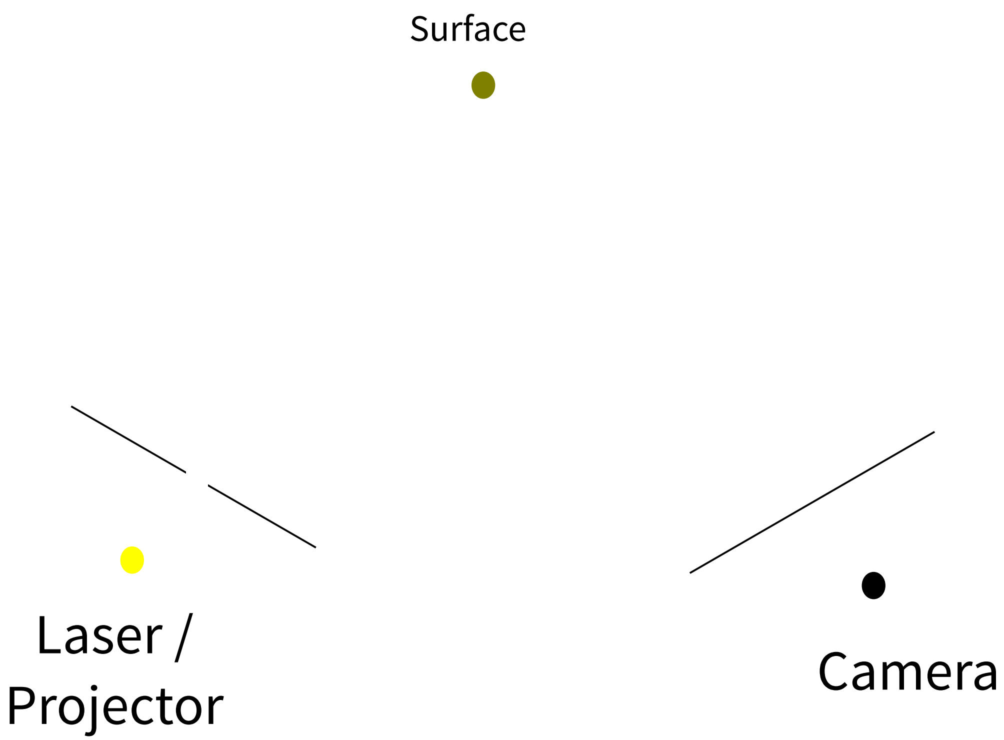
STEREO ROUNDUP

Stereo is too hard



STEREO ROUNDUP

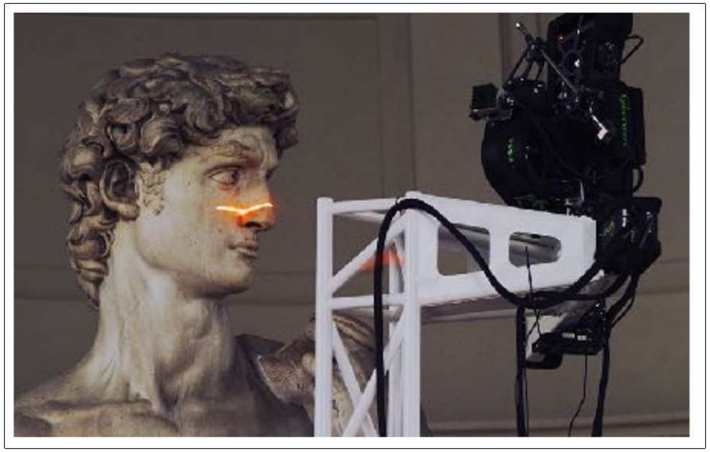
Stereo is too hard

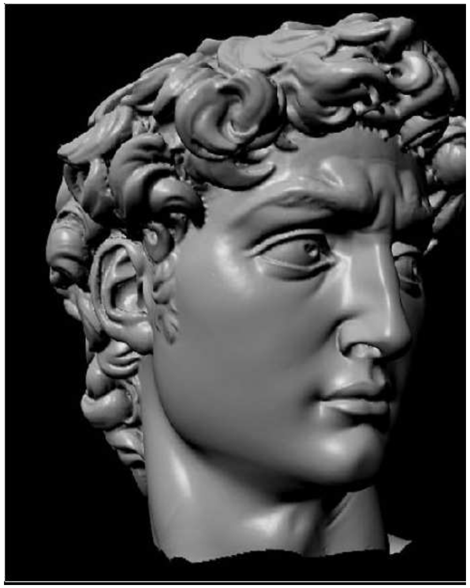


Same geometry, but we "create" the correspondences !

STEREO ROUNDUP

Stereo is too hard





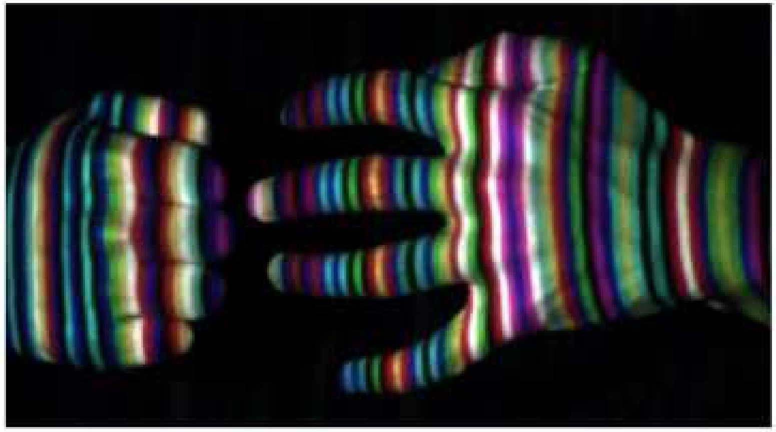
Digital Michaelangelo
Project @ Stanford

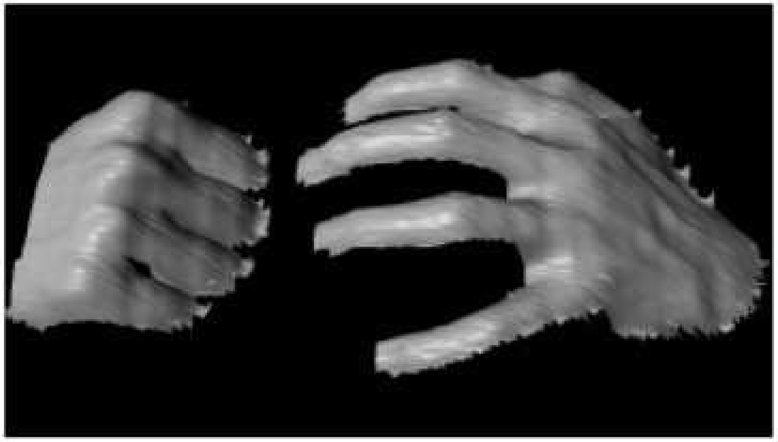
Laser Scanning

Source: Steve Seitz

STEREO ROUNDUP

Stereo is too hard





Laser Scanning too slow: Can project more general, but unique, patterns. Called "structured light" depth sensors.

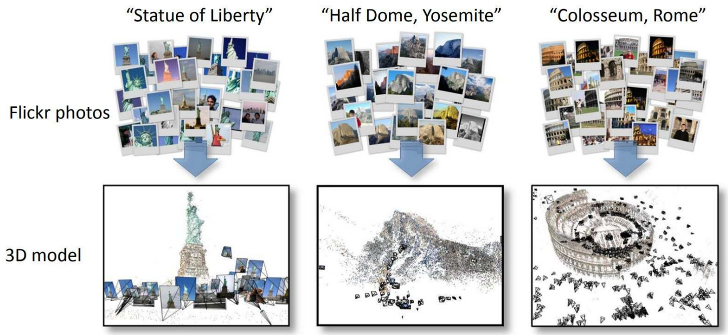
These patterns can be in infra-red

This is how the first generation of the Kinect worked.

STEREO ROUNDUP

- This is only two view stereo.

- More complicated versions include finding correspondences along multiple cameras: **Multi-view Stereo**



NEXT TIME

- General 2D Motion: Optical Flow

- Grouping Pixels

- Then onto ML